PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12Q 1/68 | A1

(11) International Publication Number: **WO 99/47706**

(43) International Publication Date: 23 September 1999 (23.09.99)

(21) International Application Number: PCT/GB99/00875

(22) International Filing Date: 19 March 1999 (19.03.99)

(30) Priority Data:
9805918.1 19 March 1998 (19.03.98) GB

(71) Applicant *(for all designated States except US)*: NYCOMED AMERSHAM PLC [GB/GB]; Amersham Place, Little Chalfont, Buckinghamshire HP7 9NA (GB).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: REEVE, Michael, Alan [GB/GB]; 118 St. Andrews Road, Henley–on–Thames, Oxfordshire RG9 1PL (GB). SCHWARZ, Terek [GB/GB]; 19 Markham Street, London, Greater London SW3 3NR (GB).

(74) Agents: PENNANT, Pyers et al.; Stevens Hewlett & Perkins, 1 Serjeants' Inn, Fleet Street, London, Greater London EC4Y 1NT (GB).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: SEQUENCING BY HYBRIDISATION

(57) Abstract

A method of analysing a nucleic acid involves the use of a mixture of labelled oligonucleotides in solution and an array of immobilised oligonucleotides. The target nucleic acid is incubated with the mixture of labelled oligonucleotides. Those labelled oligonucleotides which hybridise are recovered and incubated with the array of immobilised oligonucleotides. Sequence information is obtained by observing the location of the label on the array. The method is particularly suitable for determining differences between nucleic acids.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

SEQUENCING BY HYBRIDISATION

Introduction

There are currently three formats for sequencing by hybridisation (SBH).

Format 1 SBH [1] attaches the nucleic acid to be analysed to a solid support and then sequentially hybridises labelled oligonucleotides. Format 2 SBH [2] attaches an array of positionally encoded oligonucleotides to a solid support and then hybridises the labelled nucleic acid to be analysed to the array. Format 3 SBH [3] attaches an array of positionally encoded oligonucleotides to a solid support and then hybridises the nucleic acid to be analysed to the array in the presence of labelled oligonucleotides in free solution. A ligation reaction is then used in order to join the two oligonucleotides, giving greater specificity and information.

Format 1 SBH has been shown to work with short oligonucleotides [4]. 8 mers and even shorter oligonucleotides have been successfully employed [5]. Format 2 SBH requires the use of much longer oligonucleotides for success. 11 mer probes, or longer, are generally required. 20 mers are the norm [6], making the use of generic arrays of all **N** mers out of the question with current technology (an array of all 20 mers with the smallest pixels currently imaginable would be prohibitively large).

A difficulty with performing format 2 SBH arises because target nucleic acids often have secondary structure which sterically hinders some parts of the target from hybridising with oligonucleotides immobilised in an array. To overcome this problem it has been proposed to chop the target nucleic acid into shorter segments, e.g. of length comparable to the immobilised oligonucleotides. In practice such chopping has proved

difficult to achieve in a reliable and uniform manner. The present invention can be seen as providing an indirect way of achieving the same effect. The invention permits the advantages of both format 1 and format 2 SBH to be combined in the same method. In particular, the use of a format 2 positionally encoded array of all N mers or a subset thereof is made possible with arrayed oligonucleotides of length less than 11 mers. This method allows the rapid and facile characterisation of sequence differences between two or more nucleic acid species. The method may be used in order to determine the existence or otherwise of point mutational differences between one or more test nucleic acids and a reference nucleic acid. The method may also be used in order to characterise sequence differences arising from either small deletions or insertions.

The Invention

In one aspect the invention provides a method of analysing a target nucleic acid by the use of a mixture of labelled oligonucleotides in solution and an array of immobilised oligonucleotides, which method comprises the steps of:

a) incubating under hybridisation conditions the target nucleic acid with the mixture of labelled oligonucleotides,

b) recovering those labelled oligonucleotides that hybridised in a) with the target nucleic acid,

c) incubating under hybridisation conditions the recovered labelled oligonucleotides from b) with the array of immobilised oligonucleotides,

d) observing distribution of the labelled oligonucleotides on the array and using the information to analyse the target nucleic acid.

In another aspect the invention provides a method of determining differences between a target nucleic acid and a reference

nucleic acid, by the use of a first mixture of oligonucleotides in solution labelled with a first label, a corresponding mixture of oligonucleotides in solution labelled with a second label distinguishable from the first label, and an array of immobilised oligonucleotides, which method comprises the steps of:

a) incubating under hybridisation conditions the target nucleic acid with the first mixture of labelled oligonucleotides; and incubating under hybridisation conditions the reference nucleic acid with the second mixture of labelled oligonucleotides,

b) recovering a mixture of those first labelled oligonucleotides and those second labelled oligonucleotides that hybridised in a) with the target nucleic acid or the reference nucleic acid,

c) incubating under hybridisation conditions the recovered mixture of first labelled oligonucleotides and of second labelled oligonucleotides from b) with the array of immobilised oligonucleotides,

d) observing distribution of first labelled oligonucleotides and of second labelled oligonucleotides on the array and using the information to determine differences between the target nucleic acid and the reference nucleic acid.

Preparation of single stranded nucleic acid

The target nucleic acids may be DNA, RNA, PNA [7], other nucleic acid mimetics or mixtures thereof. They may be single stranded or double stranded; linear, circular, relaxed or supercoiled. They may be of eukaryotic, prokaryotic or viral or archeabacterial origin and may range in size from oligomers to whole genomes.

The target nucleic acids are rendered single stranded. The most preferable method is to amplify the region of interest by PCR [8] and then capture one of the amplified strands using a solid support. Many methods will be obvious to those skilled in the art. The use of a

biotinylated PCR primer followed by capture with streptavidin coated magnetic beads [9] is a preferred embodiment.

The PCR may be carried out either by using conventional dNTPs or dNTP analogues that impart altered properties to the PCR product - such as reduced intramolecular secondary structure and thus improved short oligonucleotide access to PCR product in single stranded form. Example nucleotide analogues include: dITP, 7-deaza-dGTP, 7-deaza-dATP, 7-deaza-dITP, 5-hydroxymethyl-dUTP and 4-methyl-dCTP - either singly and in combination. Many other analogues will be obvious to those skilled in the art. Some of these analogues may require the use of lower PCR annealing temperatures and / or longer PCR extension times for optimal incorporation.

The method of the invention involves use of a mixture of labelled oligonucleotides in solution. This is preferably a mixture of all or a subset of N mers where N is from 5 to 10, preferably 8 or 9. The labelling moieties may be detected by means of fluorescence (emission, lifetime or polarisation), absorption, colour, chemiluminescence, enzymatic activity, radioactive emission, mass spectroscopy or refractive index effects (e.g. surface plasmon resonance).

The N mers in solution may DNA, RNA, PNA, other nucleic acid mimetics or mixtures thereof. They may be single stranded or partially double stranded. The N mers may also contain bases such as 5-nitroindole, 3-nitropyrrole or inosine that pair with all four usual DNA bases - improving the hybridisation properties of the N mers without increasing the nucleic acid sequence complexity. The N mers may likewise contain bases such as 2-aminopurine and 5-methylcytosine that again improve the hybridisation properties without increasing the nucleic acid sequence complexity.

Structures that can only (or preferentially) form A helices are of particular interest as conditions may be found (e.g. R-loop conditions)

where the N mer / PCR product complexes are more stable than the internal secondary structure within the PCR product.

The N mers could also be molecular beacon [10] type 'panhandle' structures with stems comprising 5-nitroindole, 3-nitropyrrole, inosine, isodC:isodG [11], dk:dX [12] or dk:dp [13] hairpins. Other such structures will be obvious to those skilled in the art.

The method of the invention also involves use of an array of immobilised oligonucleotides. Each oligonucleotide is immobilised at a spaced location on a surface of a support. The array is preferably of all possible N mer sequences or a subset thereof where N is preferably from 5 to 10, particularly 8 or 9.

The array elements may be DNA, RNA, PNA, other nucleic acid mimetics or mixtures thereof. They may be single stranded or partially double stranded. The array elements may also contain bases such as 5-nitroindole, 3-nitropyrrole or inosine that pair with all four DNA bases - improving the hybridisation properties of the array without increasing its nucleic acid sequence complexity. The array elements may likewise contain bases such as 2-aminopurine and 5-methylcytosine that again improve the hybridisation properties of the array without increasing its nucleic acid sequence complexity.

Arrays may be employed on glass, plastic, silicon, supported membrane and supported gel substrates. A given substrate may have one or more test site arrays for use with the invention.

In step a) of the method, the target nucleic acid is incubated under hybridisation conditions with the mixture of labelled oligonucleotides. In step b), those labelled oligonucleotides that hybridised in a) with the target oligonucleotide are recovered. Where the target nucleic acid has been immobilised on magnetic beads as discussed above, the captured oligonucleotides may readily be recovered by denaturation and removal of the magnetic beads.

In a preferred aspect, the method of the invention may be performed to determine differences between a target nucleic acid and a reference nucleic acid. In this case, the reference nucleic acid is incubated under hybridisation conditions with a second mixture of labelled oligonucleotides, and those members of that mixture that hybridised with the reference nucleic acid are recovered. The first mixture of labelled oligonucleotides in solution is distinguishable from the second mixture of labelled oligonucleotides in solution. For example, the labels used may be fluorescent dyes having different fluorescence characteristics. The labels are herein called label 1 and label 2. Preferably the two sets of captured oligonucleotides are mixed.

In step c) the recovered mixture of labelled oligonucleotides is incubated under hybridisation conditions with the array of immobilised oligonucleotides.

Upon hybridisation to the array, captured oligonucleotides in the test and reference nucleic acids of N bases complementary to array sequences will display the normal ratio of label 1 to label 2 upon detection where the test and reference nucleic acid have the same sequence - i.e. in the majority of cases.

Upon hybridisation to the array, captured oligonucleotides in the test and reference nucleic acids of N bases complementary to array sequences will display an altered normal ratio of label 1 to label 2 upon detection where the test and reference nucleic acid have different sequence - i.e. in the vicinity of a mutation.

Difference characterisation

By observing the sequences of array elements where the label 1 to label 2 ratio is different from the majority of hybridisation events and by observing which of the two labelled moieties dominates at each such complementary array element (of known sequence), one may deduce

the sequence at and around any difference between the two nucleic acid species. In the simple case of a point mutational difference between the test and reference nucleic acid with an array of all possible N-mers, a region of 2N-1 bases will be characterised (the reference / mutated base and the N-1 bases to either side of this).

Advantages of the current invention

A particular problem that is overcome in this approach where part of the amplified single stranded region of interest has significant internal secondary structure. This situation will deny access from short oligonucleotides in solution (or as part of a positionally encoded array on a solid support). It is essentially for this reason that success has not been achieved for format 2 SBH with arrayed oligonucleotides shorter than 11 mers (arrays of 20 mers are generally used). In this invention, nucleotide analogues may be used - either in the PCR reaction or in the solution oligonucleotides or in the arrayed oligonucleotides - in order to circumvent problems with PCR product secondary structure.

This method has the advantage that by detecting perturbations in the ratio between the labelling moieties upon detection, all hybridisation events are internally controlled for their absolute hybridisation intensities - a significant improvement over other SBH methods.
Not only is information given that a difference exists between the two nucleic acid species but also the exact nature of the difference and the local sequence around this difference can be determined.

If four colour detection is implemented, the mutational event could be sequenced on both strands simultaneously - greatly improving the accuracy of an already very information-rich method.

The method does not use enzymes for the recognition of sequence differences. The method thus provides a more robust and reliable way to characterise nucleic acid sequence differences.

In addition to the above, a single array of, for example, all possible N-mers or a subset thereof, can be employed for the analysis of any nucleic acid system. Unlike other methods for sequence characterisation with arrays [6], a distinct sequence array does not need to be fabricated anew for every nucleic acid system that is to be characterised.

Unlike methods such as SSCP [14], where the optimal size for a PCR product is around 200 bp, this method allows the user to 'walk' along a genomic region of interest in much larger steps - 1-10 kb fragments would probably be about optimum for this method.

This method allows for highly parallel analysis where the shorter labelled oligonucleotides allow better mismatch discrimination. Repeated cycles of N mer capture and denaturation can be used to improve the final detection signal. Optimal chemical intermediates can selectively overcome, secondary structure. Incomplete arrays of (optimised) longer probes could be used with appropriate sequence reconstruction algorithms. Solution hybridisation to long probes and oligonucleotide hybridisation to the array should also be faster.

EXAMPLE

Step 1

All oligonucleotides were synthesised by MWG Biotech Ltd, Milton Keynes, UK).

The two target sequences comprise 45 bases spanning the human HbA and HbS beta globin gene sequences respectively. These two sequences differ at a single base position resulting in a change in codon 6 from GAG to GTG. The resulting amino acid substitution from Glu to Val in the beta globin gene is responsible for sickle cell disease. Octamer oligonucleotides corresponding to the region immediately surrounding the mutation, sequences shared by both the HbA and HbS, and sequences

mismatched by one base to one or both templates were synthesised with a terminal phosphorothioate, attached to the oligonucleotide via a $(C_{18})_3$ linker, for spotting in an array. Octamers complementary to HbA (Cy3 labelled) and HbS (Cy5 labelled) were synthesised for the solution hybridisation. The sequences are set out below.

Step 2

Synthetic template (0.1uM HbS or HbA) was subjected to a polymerase chain reaction (PCR) in the presence of 0.4uM each of the forward and biotinylated reverse sequencing primers, 250uM deoxyribonucleotides(Amersham Pharmacia Biotech), 1x PCR buffer (Amersham-Pharmacia Biotech) and 2Units of Taq DNA polymerase (Amersham-Pharmacia Biotech) in a total volume of 100µl. The thermal cycling was carried out on an MJ Research DNA engine (MJ Research inc. Watertown MA USA) for 25 cycles of: 95°C, 1min; 50°C, 30 sec; 72°C 30 sec.

Step 3

Streptavidin coated paramagnetic beads (Advanced Biotechnologies Ltd, UK) were magnetically captured in a brown 1.5ml Eppendorff tube (5mg/tube), then washed twice in 5xSSC 5% Tween-20. Following the final capture, the beads were resuspended in 0.5ml of 5xSSC, 5% Tween-20. A 100µl aliquot of PCR product from Step 2 was added to an equal volume of the bead suspension. The mixture was incubated at room temperature for 2h, with the beads maintained in suspension by continuous mixing.

Following template binding, the beads were washed three times for 10 min. with 5xSSC, 5% Tween-20 at room temperature. The beads were resuspended in 100µl of 0.05M NaOH, 1% (v/v) Tween-20 and allowed to stand for 2min. The beads were then captured and

resuspended in a second 100μl aliquot of 0.05M NaOH, 1% (v/v) Tween-20. After a further 2min. at room temperature, the beads were captured and the denaturing solution aspirated. The single stranded template bound to the beads was washed four times for 10 min. with 5xSSC, 5% Tween-20. The beads were finally resuspended in 100μl 5xSSC, 5% Tween-20.

Step 4

Hybridisation of the oligonucleotide probe mixture to bead-bound single stranded template was performed overnight at 4°C. A 50μl volume of beads was magnetically captured and the supernatant aspirated. To the captured beads 125μl of 10xSSC, 10%(v/v) Tween-20 were added and mixed with 125μl of an oligonucleotide probe mixture, in water, containing 0.4μM each labelled oligonucleotide. The beads were maintained in suspension by continuous mixing on a rolling mixer.

Following hybridisation, the beads were washed 3 times at 4°C for 20min. in 5xSSC, 5% Tween-20. After the final wash the beads were resuspended in 20μl of water and the suspension heated to 90°C for 3min. The tubes were then quickly immersed in ice water and the beads separated on an ice-cold capture magnet. An aliquot of the eluted probe was mixed with an equal volume of 10xSSC, 10% Tween-20 and used directly in a hybridisation on glass slides.

Step 5

Untreated glass slides (Erie Scientific, Portsmouth New Hampshire, USA, Cat No. 2959F*) were soaked in 5%(v/v) (3-mercaptopropyl)trimethoxysilane (Aldrich Chemical Co., Poole, Dorset, UK) in dry toluene for 6 hours. The slides were washed with dry toluene followed by ethanol. The slides were then soaked overnight in a 6.66 g/l isopropanol solution of 2,2'-dipyridyl disulfide (Sigma Chem. Co. UK). The slides were finally washed three times with isopropanol and air-dried.

Step 6

The microtitre plates, containing oligonucleotides to be arrayed, were prepared by mixing in each well 5μl of an oligonucleotide solution (20 pmol/ l), 10 μl of 50%(v/v) aqueous ethyleneglycol and 5 l of imidazole buffer pH4. Glass slides prepared in Step 5 were spotted with the oligonucleotide solutions using a Molecular Dynamics spotter set in normal mode at a humidity of 45-47% and a temperature of 20°C. Once spotted, the slides were kept over night at room temperature in a humidified chamber. The slides were then washed with water, rinsed with isopropanol and allowed to dry.

Step 7

Probe solutions (10μl aliquots) from Step 4 were applied to the arrays prepared in Step 6 and covered with a coverslip. The slides, kept above water in a sealed box, were heated to 60°C in an Amersham Pharmacia Biotech hybridisation oven, kept at this temperature for 1 hour, then cooled in the oven to room temperature over a period of 3 hours, followed by 2 hours at room temperature. The slides were washed twice for 3 minutes at 4°C with 5xSSC, 5% Tween-20. The washed slides were scanned for the Cy3 and Cy5 fluorescence signals using a Molecular Dynamics generation III micro-array scanner with default settings, 695V for the green laser, 750 V for the red laser. The scans were analysed using Molecular Dynamics' Image Quant and Microsoft Excel software.

Results

HbS template hybridised with Cy5 and Cy3 labelled oligonucleotide mixture.

HbS template was subjected to a hybridisation described in Step 4 and the resulting eluate of oligonucleotides was hybridised to the

array of ocatmers described in Step 7. The resulting hybridisation pattern showed all eight octamers tiled across the HbS polymorphism had hybridised and produced a signal. Hybridisation signal with both Cy5 and Cy3 was also observed for the oligonucleotides that are common to both templates. Negligible signal was observed for any of the mismatched oligonucleotides, thus demonstrating the capture and hybridisation of template specific oligonucleotides. A number of array elements failed to produce a signal when hybridisation was performed directly with the Cy3 and Cy5 oligonucleotides and were excluded from the analysis as void elements.

HbA template hybridised with Cy5 and Cy3 labelled oligonucleotide mixture

The HbA template was subjected to analysis as described for the HbS template above. HbA specific Cy3 fluorescece was observed on the array. Five of the Eight HbA specific elements of the array gave a positive signal. The negative elements that would have been expected to give a signal when hybridised to a matched probe were also negative when the fluorescent oligonucleotides were added directly to the array. This finding confirmed that the absence of signal at these points was the result of spot failure, rather than absence of the appropriate fluorescent oligonucleotide. Functional array elements representing shared sequence showed up with both Cy3 and Cy5 fluorscence signals. HbS specific array elements and mismatched oligonucleotides gave either faint signal or no signal at all

HbA and HbS templates hybridised with Cy5 and Cy3 labelled oligonucleotides

When both HbA and HbS templates were present all Functional elements of the array gave a fluorescent signal for matched oligonucleotides; with the HbS and HbA elements of the array producing

Cy5 and Cy3 signals respectively. Common sequences produced a signal with both Cy3 and Cy5 oligonucleotides. Mismatched oligonucleotides showed only faint singal or no signal in both Cy3 and Cy5 channels.

Templates

HbA SEQUENCE

GTTTTCCCAGTCACGACG
ACCATGGTGCACCTGACTCCT**GAG**GAGA
AGTCTGCCGTTACTGCC ATGGTCATAG
CTGTTTCCT

HbS SEQUENCE

GTTTTCCCAGTCACGACG
ACCATGGTGCACCTGACTCCT**GTG**GAGA
AGTCTGCCGTTACTGCC ATGGTCATAG
CTGTTTCCT

Underlined sequences represent a primer binding sequence, or its complement, that has been appended to the Hb sequence. The boxed nucleotides in bold type highlight the single base polymorphism of codon 6.

Forward primer

GTT TTC CCA GTC ACG ACG

Reverse primer

5' Biotin-AGG AAA CAG CTA TGA CCA T

Reverse sequencing primer is biotinylated at the 5' end to permit capture on streptavidin coated surfaces.

Oligonucleotides for preparing the array

P(s)=phosporothioate, X=C18 linker.

HbA specific oligonucleotides

5'-p(s)-XXX-**T**CAGGAGT-3'
5'-p(s)-XXX-C**T**CAGGAG-3'
5'-p(s)-XXX-CC**T**CAGGA-3'
5'-p(s)-XXX-TCC**T**CAGG-3'
5'-p(s)-XXX-CTCC**T**CAG-3'
5'-p(s)-XXX-TCTCC**T**CA-3'
5'-p(s)-XXX-TTCTCC**T**C-3'
5'-p(s)-XXX-CTTCTCC**T**-3'

HbS specific oligonucleotides
5'-p(s)-XXX-**A**CAGGAGT-3'
5'-p(s)-XXX-C**A**CAGGAG-3'
5'-p(s)-XXX-CC**A**CAGGA-3'
5'-p(s)-XXX-TCC**A**CAGG-3'
5'-p(s)-XXX-CTCC**A**CAG-3'
5'-p(s)-XXX-TCTCC**A**CA-3'
5'-p(s)-XXX-TTCTCC**A**C-3'
5'-p(s)-XXX-CTTCTCC**A**-3'

Oligonucleotides common to both templates
5'-p(s)-XXX-CGTCGTGA-3'
5'-p(s)-XXX-GTCGTGAC-3'
5'-p(s)-XXX-TCGTGACT-3'
5'-p(s)-XXX-CGTGACTG-3'
5'-p(s)-XXX-GTGACTGG-3'
5'-p(s)-XXX-ATGACCAT-3'
5'-p(s)-XXX-TATGACCA-3'
5'-p(s)-XXX-CTATGACC-3'
5'-p(s)-XXX-ACTATGAC-3'
5'-p(s)-XXX-CACTATGA-3'

Oligonucleotides with one base mismatch to either HbA, HbS, or both.

5'-p(s)-XXX-CTCATCAG-3'
5'-p(s)-XXX-CTCGTCAG-3'
5'-p(s)-XXX-CTCTTCAG-3'
5'-p(s)-XXX-CTCAACAG-3'
5'-p(s)-XXX-CTCGACAG-3'
5'-p(s)-XXX-CTCTACAG-3'
5'-p(s)-XXX-TCGAGACT-3'
5'-p(s)-XXX-TCGCGACT-3'
5'-p(s)-XXX-TCGGGACT-3'
5'-p(s)-XXX-CTAAGACC-3'
5'-p(s)-XXX-CTACGACC-3'
5'-p(s)-XXX-CTAGGACC-3'

Fluorescently labelled capture oligonucleotides.

5'Cy3-ACTCCTGA3'
5'Cy3-CTCCTGAG3'
5'Cy3-TCCTGAGG3'
5'Cy3-CCTGAGGA3'
5'Cy3-CTGAGGAG3'
5'Cy3-TGAGGAGA3'
5'Cy3-GAGGAGAA3'
5'Cy3-AGGAGAAG3'
5'Cy3-TCACGACG3'
5'Cy3-GTCACGAC3'
5'Cy3-AGTCACGA3'
5'Cy3-CAGTCACG3'
5'Cy3-CCAGTCAC3'
5'Cy3-ATGGTCAT3'
5'Cy3-TGGTCATA3'
5'Cy3-GGTCATAG3'
5'Cy3-GTCATAGCT3'
5'Cy3-TCATAGCTG3'

5'Cy5-ACTCCTGT3'
5'Cy5-CTCCTGTG3'
5'Cy5-TCCTGTGG3'
5'Cy5-CCTGTGGA3'
5'Cy5-CTGTGGAG3'
5'Cy5-TGTGGAGA3'
5'Cy5-GTGGAGAA3'
5'Cy5-TGGAGAAG3'
5'Cy5-TCACGACG3'
5'Cy5-GTCACGAC3'
5'Cy5-AGTCACGA3'
5'Cy5-CAGTCACG3'
5'Cy5-CCAGTCAC3'
5'Cy5-ATGGTCAT3'
5'Cy5-TGGTCATA3'
5'Cy5-GGTCATAG3'
5'Cy5-GTCATAGCT3'
5'Cy5-TCATAGCTG3'

References

[1] Drmanac, R., Labat, I., Brukner, I., Crkvenjakov, R., Sequencing of megabase plus DNA by hybridization: theory of the method. Genomics 4, 114-128.

[2] Khrapko KR, Yu P, Khorlyn M, Shick W, Florentiev VL & Mirzabekov 1989, An oligonucleotide hybridization approach to DNA sequencing, FEBS Lett. 256:118-122.

[3] Hyseq Inc. ATP Project Brief 94-05-0018

[4] Strezoska Z, Paunesku T, Radosavljevic D, Labat I, Dramanac R & Crkvenjakov R 1991, DNA sequencing by hybridization: 100 bases read by a non-gel-based method, PNAS 88: 10089-10095.

[5] Drmanac R, Drmanac S, Strezoska Z, Paunesku T, Labat I, Zeremski M, Snoody J, Funkhouser WK, Koop B, Hood I & Crkvenjakov R 1993, Science 260: 1649-1652.

[6] Hacia JG, Brody LC, Chee MS, Fodor SPA & Collins FS 1996, Detection of heterozygous mutations in BRAC1 using high density arrays and two-colour fluorescence analysis. Nature Genetics 14: 441-447.

[7] Nielsen PE, Egholm M, Berg RH & Buchardt O 1991, Sequence-selective recognition of DNA by strand displacement with a thymine-substituted polyamide. Science254: 1497-1500.

[8] Saiki RK, Scharf S, Faloona F, Mullis KB, Horn GT, Erlich HA & Arnheim N 1985, Enzymatic amplification of b-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anaemia. Science 230: 1350-1354.

[9] Espelund M, Prentice RA & Jakobsen KS 1990, A simple method for generating single-stranded DNA probes labeled to high activities. Nucleic Acid Research 18: 6157-6158.

[10] Tyagi S & Kramer FR 1996, Molecular beacons: probes that fluoresce upon hybridization. Nature Biotechnology 14: 303-308.

[11] Switzer C, Moroney SE & Benner SA 1989, Enzymatic incorporation of a new base pair into DNA and RNA J. Am. Chem. Soc. 111: 8322-8323.

[12] Piccirilli JA, Krauch T, Moroney SE & Benner SA 1990, Enzymatic incorporation of a new base pair into DNA and RNA extends the genetic alphabet, Nature 343: 33-37 and Horlacher J, Hottiger M, Podust VN, Hubscher U, & Benner SA 1995, Recognition by viral and cellular DNA polymerases of nucleosides bearing bases with nonstandard hydrogen bonding patterns. PNAS 92: 6329-6333.

[13] Piccirilli JA, Krauch T, Moroney SE & Benner SA 1990, Enzymatic incorporation of a new base pair into DNA and RNA extends the genetic alphabet, Nature 343: 33-37.

[14] Orita M et al 1989, Rapid and sensitive detection of point mutations and DNA polymorphisms using the polymerase chain reaction. Genomics 5: 874-879.

CLAIMS

1. A method of analysing a target nucleic acid by the use of a mixture of labelled oligonucleotides in solution and an array of immobilised oligonucleotides, which method comprises the steps of:

a) incubating under hybridisation conditions the target nucleic acid with the mixture of labelled oligonucleotides,

b) recovering those labelled oligonucleotides that hybridised in a) with the target nucleic acid,

c) incubating under hybridisation conditions the recovered labelled oligonucleotides from b) with the array of immobilised oligonucleotides,

d) observing distribution of the labelled oligonucleotides on the array and using the information to analyse the target nucleic acid.

2. A method as claimed in claim 1, wherein the target nucleic acid used in a) is immobilised.

3. A method of determining differences between a target nucleic acid and a reference nucleic acid, by the use of a first mixture of oligonucleotides in solution labelled with a first label, a corresponding mixture of oligonucleotides in solution labelled with a second label distinguishable from the first label, and an array of immobilised oligonucleotides, which method comprises the steps of:

a) incubating under hybridisation conditions the target nucleic acid with the first mixture of labelled oligonucleotides; and incubating under hybridisation conditions the reference nucleic acid with the second mixture of labelled oligonucleotides,

b) recovering a mixture of those first labelled oligonucleotides and those second labelled oligonucleotides that hybridised in a) with the target nucleic acid or the reference nucleic acid,

c) incubating under hybridisation conditions the recovered mixture of first labelled oligonucleotides and of second labelled oligonucleotides from b) with the array of immobilised oligonucleotides,

d) observing distribution of first labelled oligonucleotides and of second labelled oligonucleotides on the array and using the information to determine differences between the target nucleic acid and the reference nucleic acid.

4. A method as claimed in claim 3, wherein the target nucleic acid used in a) is immobilised, and the reference nucleic acid used in a) is immobilised.

5. A method as claimed in any one of claims 1 to 4, wherein the or each mixture of labelled oligonucleotides used in a) is a complete set or a subset of all N-mers where N is 5 to 10.

6. A method as claimed in any one of claims 1 to 5, wherein the array of immobilised oligonucleotides used in c) is an array of a complete set or a subset of all N-mers where N is 5 to 10.

7. A method as claimed in claim 5 or claim 6, wherein N is 8 or 9.

8. A method as claimed in any one of claims 1 to 7, wherein the labelled oligonucleotides in solution are DNA, RNA, PNA, other nucleic acid mimetics or mixtures thereof, are single stranded or partially double stranded, and may contain residues of one or more nucleotide analogues.

9. A method as claimed in any one of claims 1 to 8, wherein the oligonucleotides of the array are DNA, RNA, PNA, other nucleic acid mimetics or mixtures thereof, are single stranded or partially double stranded, and may contain residues of one or more nucleotide analogues.

10. A method as claimed in any one of claims 1 to 9, wherein the target nucleic acid is generated by PCR.

11. A kit for performing the method of any one of claims 1 to 10, or comprising a supply of a mixture of labelled oligonucleotides in solution and an array of immobilised oligonucleotides.

INTERNATIONAL SEARCH REPORT

International Application No
PCT/GB 99/00875

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12Q1/68

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 95 09248 A (ARCH DEVELOPMENT CORP.) 6 April 1995 (1995-04-06) page 14, line 4 - page 15, line 36; claims 35-42	10
A	US 5 683 881 A (S.S.SKIEMA) 4 November 1997 (1997-11-04) column 6, line 40 - column 7, line 28; claims	1
A	K.R.KHRAPKO ET AL.: "A method for DNA sequencing by hybridisation with oligonucleotide matrix" DNA SEQUENCE-I. DNA SEQUENCING AND MAPPING, vol. 1, 1991, pages 375-388, XP002910053 UK the whole document	1

☐ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
25 August 1999	31/08/1999

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Luzzatto, E

1

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/GB 99/00875

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9509248 A	06-04-1995	AU 694146 B	16-07-1998
		AU 8072794 A	18-04-1995
		BR 9407712 A	12-02-1997
		CA 2172722 A	06-04-1995
		CN 1136330 A	20-11-1996
		CZ 9600905 A	16-10-1996
		EP 0723598 A	31-07-1996
		FI 961283 A	22-05-1996
		HU 75993 A	28-05-1997
		JP 9505729 T	10-06-1997
		NO 961165 A	23-05-1996
		NZ 275194 A	22-09-1997
		PL 313735 A	22-07-1996
US 5683881 A	04-11-1997	NONE	